# Object Detection and Localization on Map using Multiple Camera and Lidar Point Cloud

Pansipansi, Leonardo John* · Minseok Jang · Yonsik Lee

Kunsan National University

E-mail : 2010074@kunsan.ac.kr, {msjang, yslee}@kunsan.ac.kr

## ABSTRACT

In this paper, it leads the approach of fusing multiple RGB cameras for visual objects recognition based on deep learning with convolution neural network and 3D Light Detection and Ranging (LiDAR) to observe the environment and match into a 3D world in estimating the distance and position in a form of point cloud map. The goal of perception in multiple cameras are to extract the crucial static and dynamic objects around the autonomous vehicle, especially the blind spot which assists the AV to navigate according to the goal. Numerous cameras with object detection might tend slow-going the computer process in real-time. The computer vision convolution neural network algorithm to use for eradicating this problem use must suitable also to the capacity of the hardware. The localization of classified detected objects comes from the bases of a 3D point cloud environment. But first, the LiDAR point cloud data undergo parsing, and the used algorithm is based on the 3D Euclidean clustering method which gives an accurate on localizing the objects. We evaluated the method using our dataset that comes from VLP-16 and multiple cameras and the results show the completion of the method and multi-sensor fusion strategy.

### Keywords

Autonomous vehicle, Multi-Cameras, Point Cloud Clustering, Object Detection, Localization

## Ⅰ. Introduction

Autonomous vehicles or self-driving vehicles provide the transportation capabilities of conventional vehicles but are largely capable of perceiving the environmental and self-navigating with minimal or no human intervention. The ideal for autonomous vehicles in the future can go anywhere, convenient to ride, and safe. In this project, we researched perception of the environment and localize. The sensor being used in this project are VLP-16 and multiple cameras which are fused to perceive the environment. Lidar sensor sparse point cloud data which determine the distance of all object in the environment and cameras can perceive the object using deep learning method. Therefore this combination of the sensor is adequate for the autonomous car. The fusion of multi-camera and lidar sensors has the advantage, it can be classified as dynamic or stationary as well as by type (e.g. vehicle, pedestrian, or other). The ability of sensors that can detect objects using deep learning methods and identify distance estimation can utilize in terms of avoiding the collision.

## Ⅱ. Implementation Method

In this proposed paper we will discuss the important method of object detection and localization using multiple cameras and lidar point cloud for autonomous vehicles. The multiple sensors of this project will go through the fusion process shown in figure 1 below. The fusion of sensors perceives the situation of the surrounding by obtaining the data from multiple cameras and lidar point cloud.


Figure 1. Multi sensor lidar and cameras.

* corresponding author

The first method is lidar camera calibration which involves two parts in the process intrinsic and extrinsic parameter calibration. Lidar camera calibration estimates a rigid transformation matrix that establishes the correspondences between the points in the 3D lidar plane and the pixels in the image plane.

Extrinsic lidar and camera calibration sensors generally use calibration objects, such as planar boards with chessboard patterns, in the captured scene. The projection pose of sensors camera and lidar (Figure 2) information allows us to compute the pose of the board reference frame relative to the camera with the use of the Perspective-n-Point (PnP) algorithm in our camera and lidar calibration. The result is shown in figure 3.

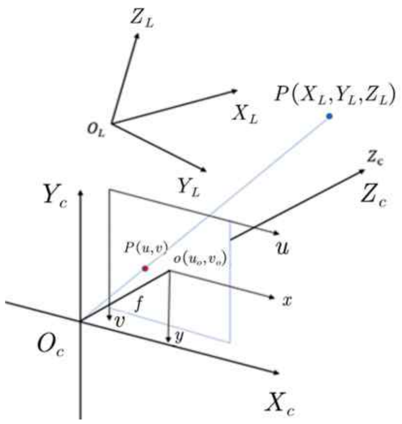



Figure 2. Projection odometry of lidar and camera are in the same direction.

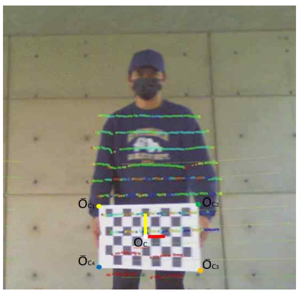
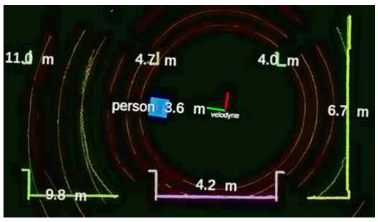


Figure 3. Point cloud overlay image

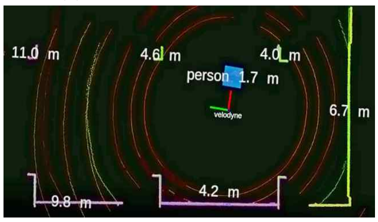
The next method is to identify an object detection algorithm that would fit the capacity of the computer in computation data since we are using multiple cameras and lidar.


Figure 4. Real-time object detection

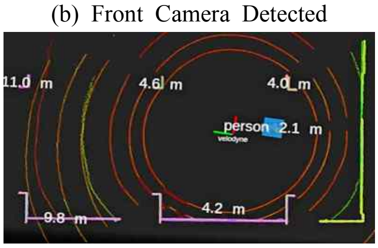
In our sensor VLP-16 considering common lidar data has a large amount of information, there are often more than 100,000 points in one frame, and it can acquire the data at a fast speed. Therefore to deal with this massive data we come up to use euclidean lidar clustering algorithm but before that, we need to filter the ground points using the RANSAC algorithm. Localization and distance estimation is the last part of this method. The fusion of these multiple cameras and lidar data will give back accurate data information. A point of the lidar point cloud is defined by a distance and two angles. To represent the two angles it uses azimuth(Θ) and polar angle(ϕ) convention. Thus a point is defined by (r,Θ,ϕ) or a 3D coordinate.




(a) Left Camera Detected




(b) Front Camera Detected




(c) Right Camera Detected

Figure 5. Object detection and localization

Table 1. Lidar-Camera distance estimation

| | Classfied Object | Distance Estimation |
|---|---|---|
| Left Camera | Pedestrian | 3.6 m |
| Front Camera | Pedestrian | 1.7 m |
| Right Camera | Pedestrian | 2.1 m |

## Ⅲ. Conclusion

The algorithm used in this project is for a real-time purpose which uses also 4 overall sensors. The result of this project uses multiple cameras and a 360 degrees lidar. the objects who are roaming around the vehicle can be identified and localize with accurate distance estimation. It can be classified what object is on the right, left, and front sides of an autonomous vehicle. This can help to eradicate the accidents of blind spot areas on the road, which sensor can detect the incoming nearest object towards the autonomous vehicle.

## Acknowledgement

This research was funded and conducted under the Competency Development Program for Industry Specialist of Korean Ministry of Trade, Industry and Energy (MOTIE), operated by Korea Institute for Advancement of Technology (KIAT) (N0002428, HRD Program for Future Car), a grant (21RITD-C161698-01) from Regional Innovation Technology Development Program funded by Ministry of Land, Infrastructure and Transport of Korean government and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2021R1F1A1047768).

## References

[1] Pan Wei, Lucas Cagle, Tasmia Reza, John Ball and James Gafford., "Lidar and Camera Detection Fusion in a Real-Time Industrial Multi-Sensor Collision Avoidance System.", MDPI journal Electronics, 7(6), 84, May, 2018

[2] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[3] Anandakumar M. Ramiya, Rama Rao Nidamanuri, Ramakrishan Krishnan, Segmentation based building detection approach from LiDAR point cloud, The Egyptian Journal of Remote Sensing and Space Science, Volume 20, Issue 1, 2017,

[4] S. Kato et al., “Autoware on Board: Enabling Autonomous Vehicles with Embedded Systems,” 2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS), 2018, pp. 287-296, doi: 10.1109/ICCPS.2018.00035.

[8] Yeong, D. J.; Velasco-Hernandez, G.; Barry, J.; Walsh, J. Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review. Sensors 2021, 21, 2140